The Making of Meaning in Societies: Semiotic & Information-Theoretic Background to the Evolution of Communication

Chrystopher L. Nehaniv
Interactive Systems Engineering
University of Hertfordshire
College Lane
Hatfield Herts AL10 9AB
United Kingdom
C.L.Nehaniv@herts.ac.uk

Abstract

We examine the notions of meaning and information for animals or agents engaged in interaction games. Concepts from cognitive ethology, linguistics, semiotics, and evolution are surveyed. Innateness, individual learning, and social aspects (social learning and cultural transmission) of the evolution of communication are treated. Studies on animals and agents showing degrees of communication are analyzed with an eye to describing what aspects of communication actually are demonstrated, or also in the case of many simulation studies, are built-in to the system at the outset. In particular, predication and constituent structure (subcategorization) have so far never been shown to emerge in robotic or software systems.

1 Introduction

Meaning in real human societies is socially constructed (Bruner (1991)), yet this depends also an the individual member of society's participation. The making of meaning in society emerges from the interaction of many participants as they communicate to one another about the world in which they are situated. Obviously the participants have particular biological capacities necessary for a construction of meaning, but the degree to which innate mechanisms as opposed to learning or cultural mechanisms are involved is the subject of much debate, especially in the case of human language acquisition. For other animals and for software and robotic agents, evolving or designing communication systems present similar issues. The substrate upon which communication relies can be compared and contrasted to the human case, and the insights should be useful in several areas: (1) understanding human communication and language situated in the context of a general biological background, (2) identification and description of characteristics, properties, and mechanisms sufficient for the support of communication systems of various kinds in animals, (3) the design and construction of mechanisms to support communication and language-like phenomena in artificial systems.

Semiotics provides an insightful approach to understanding meaning in terms of a *relational* (rather than a naive *mapping*) framework (Peirce (1995); Goguen (1999)). In particular, a *sign* or signal is related to a *signified* via an *interpretant*, the situated linkage between the two, depending on participant in the particular act of *semiosis*. The segregration of the *sign* and also of the *signified* from the background of the environment are not a priori given, nor need they coincide for different participants in acts of semiosis. (Although this theory of meaning seems simple enough, it is much more complex than what one usually sees in agent studies of the "evolution of language" or "evolution of communication", which assume a (generally fixed) set of possible referents and (generally fixed) set or alphabet of signs, both available to all agents at the outset.) The legs of the semiotic triad (sign, signified, interpretant) all vary with the particular agent in question. Thus the study of meaning is inherently an agent-oriented research area, rather than a third-person God's eyeview Platonic world of absolutes.

An information-theoretic approach can be used to study the evolution of channels of meaning in a community of agents (Nehaniv (1999)). At a fundamental level, modes of sensing and actuating afford an agent its access to potentially meaningful information – *meaningful information for a particular agent is information that is, in a statistical sense,* useful *for the agent in achieving its goals.* In addition to the interaction channels, internal structure and history of the agent also plays a critical role in facilitating the use of meaningful information to achieve goals. Applying Shannon information theory (Shannon and Weaver (1963)) to information in channels that are meaningful in this sense allows one to develop an agent-based theory of meaning as an extension of the mathematical theory of communication.

The foregoing remarks already apply to a single agent or animal interacting with its *Umwelt*, the ethologist's term for the environment in which it is embodied and embedded. Moreover, for social animals, and for socially intelligent agents, meaning (in the sense just outlined) emerges from the interaction of semiotically active agents comprising the society. Which goals are desirable for an agent depends on its nature, but also the culture in which it developed, channels of information that are meaningful for attaining these goals are in part determined by design (evolutionary or intentional) and in part by the history of interaction with others. The segregation of signs and signified from a morass of environmental stimuli to comprise legs of the semiotic triad (within components of a system of signs) depends also on embodiment, society and history of interaction.

We argue that useful models of the evolution of communication must take into account the principles described here, and that other current models of fatally flawed methodologically or at best incomplete. Indeed many published results in the evolution of communication can be shown to be consequences of random statistical sampling errors leading to convergence of (naive) "communication systems" in which the potential signs and their referents were circumscribed by experimenters at the outset and in which a (naive) notion of semantics constrained the nature of the possible systems which could evolve — only in a manner that would seem to confirm the preconceptions of the experimenters. Similar remarks apply to the "emergence of syntax" in which constituent structure (essentially context-free language formalism) has been built-in at the outset (e.g. as "slots" in semantic processing).

Principles for socially and semiotically realistic studies of the evolution of meaning could be carried out will be described, making reference to some fundamental studies on the grounding of communication (Wittgenstein (1958, 1968); Billard and Dautenhahn (1999); Dautenhahn (1995); Nehaniv et al. (1999)).

We will throw out several assumptions that are made with traditional denotational semantics, by making contrasting assertions:

1. **No Agent, No Meaning** It will not be possible to have a God's eye view notion of meaning. A signal or message can only be meaningful for particular individuals involved in particular interactions with their environment or with each other.

2. **No Privileged Meanings** We will not assume there is a special set of concepts and predicates characterizing the set of what it is possible for the agent to mean. This is for instance a rejection of a Platonic realm of forms, which the real world is only a shadow of, etc. Moroever, this entails that we may not assume a priori that certain categories (classes of objects, attributes, abstractions, etc.) exist outside agents and their interaction. The existence of such categories must always be grounded in the particular agent's internal architecture, e.g. the state of its neurons, etc., as they relate to its previous experience and interaction with others and the world.

3. **No Privileged Signals** We will not assume that there are specific, atomic symbols or classes of symbols to which all agents may in principle have access. The sensory and actuator characteristics, as well as learning and experience, conspire to determine what type of event constitutes a signal for the particular agent in question.

4. **No Privileged Mapping** Agents may have incomplete knowledge of symbols and referents, actions, and meanings that might be communicated. Moreover no particular mapping of signals to signifieds is the privileged correct mapping. Agents may have different and conflicting mappings (or relations) with different domains and ranges.

Thus each vertex of the semiotic triangle is subject to variation. Different agents use different interpretants (hence potentially different mappings) to relate sign and signified.

2 Semiosis: The Making of Meaning

A much less naïve theory of how meaning arises than the denotational semantics common in computer science is semiotics (Peirce (1995)), introduced by an American philospher working over a hundred years ago.

2.1 Semiotic Triangles

Semiotics acknowledges the situated nature of the making of meaning. The connection between a *sign* and what it signifies (the *signified*) is mediated by an *interpretant* (the relation between them). The naturalness of this relationship has degrees: A sign may be iconic (sensorially indicative of the signified), indexical (indicative but not representing the signified in a way closely matching the perceptual stimulus the signified would produce), or symbolic (arbitrarily associated to the signified). Examples of iconic signs include threatening displays in animals, indexical signs include the intention movements of animals or a hole in a wall indicating that a bullet passed through it, and symbols include arbitrary phonemic strings of spoken human language.

By making the interpretant explicit, Peirce made clear that the relationship between sign and signified is not a static one, it can vary with the agent involved, between agents, and with context. Sign, signified, and interpretant are vertices of a triangle on which each process of making meaning is based. Such a process is called *semiosis*.

The above rejection of the assumptions of denotational semantics and similar systems amounts to recognizing that each aspect of semiosis — sign, signified, and interpretant — is thus agent-particular rather than part of some external structure.

3 Meaning is (Statistically) Useful Information in Channels of Sensing and Actuating

We now relate the semiotic notion of meaning to its situated and embodied contexts in human, animal, and other agent systems.

3.1 Wittgenstein and Meaning in Use

Denotation of words may be relatively unambiguous for proper names, but general concrete terms, actions, attributes, and relationships correspond to no particular entities in the physical world.

Wittgenstein pointed out that to know the meaning of a word one must know the function of the word in the contexts in which it is used. Generalizing from his insights, we shall insist that the meaning of signals can be and should only be defined in terms of their usage in *interaction games* (Nehaniv (1999)). Animals do not evolve signal systems for the purpose of making 'true' assertions about the physical world. They are not concerned with truth, but rather with survival in the natural world. If they can use signals to manipulate the world and gain useful information about it, then this is meaningful for them and can motivate natural selective pressure.

Meaning is understood here as (1) information in interaction games between an agent and its environment or between agents mediated by the environment and in all cases by the sensors or actuators of the agents, and as (2) *useful* (in a probabilistic sense taking into account the costs and benefits of sensing and actuating) for satisfying homeostatic or other drives, needs, goals, or intentions. (see also Nehaniv (1999), Nehaniv et al. (1999)).

3.2 Private Meaning

The definition of meaning above is made with reference to a particular agent (or possibly a community), since the notion of "useful" requires this and since the notion depends also on the particular sensing and actuating capacity of the agent. Thus information that is meaning for one agent may be imperceptible or meaningless noise to another. Moreover, the internal state and structure of the agent is crucial to whether information might be useful to it. This is closely connected with whether the agent can *use* the information to modify its *expectations* (e.g. predictive scenarios) of what is likely to happen and thus modify its own future actions in light of these. (Also compare the discussion of Smith (1996) below).

4 Evolution of Communication

Darwin (1872) recognized the importance of the expression of emotion in an animal as cues by which others can judge aspects of its internal state, and thus its likely future behaviour. Cues, communicative signalling, and misinformation are distinguished in the literature on animal communication and information-theoretic properties are related via cost-benefit trade-offs to the study of the evolution of communication.

4.1 Definitions of Communication

(Bradbury and Vehrencamp (1998)) define communication as follows: "The process of communication involves two individuals, a sender and a receiver. The sender produces a signal which conveys information. The signal is transmitted through the environment and is detected by the receiver. The receiver uses the information to help make a decision about how it should respond. The receiver's response affects the fitness of the sender as well as its own. In true communication, both sender and receiver benefit (on average) from the information exchange."

Stimuli produced by an animal but not benefitting it perceived by others are called *cues*. If the production of the signal does not on average benefit the receiver, then this is called *misinformation*. Examples include the mimicry of one species' sexual pheromones by another in order to attract the former as prey, the use of fishing bait, but also camouflage and disruptive displays in animals (e.g. cephalopods Moynihan (1985); Hanlon and Messenger (1996)). (Misinformation is sometimes called "dishonest communication", but we avoid this term in that it leads to presuppositions that the receiver is capable of holding a false belief or that the emitter intends the receiver to form a false belief, etc.) Signals may be very extended in temporal extent, *states* (e.g. permanent coloration markings on the body, fixed body scents), or *events* of more limited scope (alarm calls, a display of out-spread tail feathers, aggressive posturing and coloration, etc.).

Many definitions, not requiring benefit on average to the recipient, of a signal occur in ethology:

"Communication is the phenomenon of one organism producing a signal that, when responded to by another organism, confers some advantage (or the statistical probability of it) to the signaler or his group." (Burghardt (1977))

This definition is used by MacLennan (1992) in a synthetic computational ethological implementation. Populations of "simorgs" (essentially look-up tables giving functions from global environment and local environment to either emissions and actions) are subjected to digital evolution in which they are rewarded for actions matching the local environment of the last emitter. Comparing evolution (using a steady-state genetic algorithm) of such simorgs to others for which communication was not permitted, MacLennan showed that Burghardt's definition is satisfied.

4.2 Expectation, Prediction, and Action

(Smith (1977, 1996)) considers that an animal's basic cognitive activity is characterized by "a continuous cycle of generating and testing expectations that are incorporated into predicative scenarios". Expanding this: The animal is seeking or extracting information from various sources, in various circumstances; it compares this information with information it has previously stored; and it makes and updates predictions, selects among them and generates new ones. This is a continuous process, in which information is used to produce expectations. Signals from other animals are an important source of such information. The information and predictions of an individual are largely "private", i.e. not visible to others, but may be made public by specialized behaviour called signalling, e.g. information about what the individual is likely to do next. Signaling behaviour can influence the recipient's behaviour in a manner that is useful to the sender. The behaviour of populations that signal will co-evolve with the dispositions of how recipients respond whether the recipients be in the sender's own population or another allospecific group. Formalization of signal repertoires, specialization of displays, modes of varying display form, modes of combining displays, and formalization of interactions will all be driven by the costs and benefits of signalling behaviour, and are especially likely to have effects on recipient expectation of social events (Smith (1996)). Moreover, Smith emphasizes that formalization of signalling interaction enables each participant to elicit signalling responses within formal (and thus more predictable) constraints. Here we have the evolution of interaction games (including the signalling, sensory, and processing apparatus) leading to the formalization of signalling exchanges.

The communication behaviour here arises in evolving populations engaged in social and nonsocial interaction. The nonsocial components have to do with manipulation of the environment, of predator, and of prey; while the social component can be largely (but perhaps not completely) identified with intraspecific interaction (territoriality, mate attraction, etc.). Cues such as direction of eye gaze and joint attention or signals of intention movements may be interpretable across several species, and might be considered candidates for interspecies communication (subject to further conditions of the various competing definitions).

4.3 Communicative Systems

Animal communication thus is clearly subject to inherited genetic and developmental factors. Innate signalling systems might be refined by experience, e.g. young Vervet monkeys may make inappropriate alarm calls, ignored by adults, before they can distinguish harmless birds from aerial predators, (Seyferth and Cheney (1986)). Chomsky (1968, 1975), Pinker and Bloom (1990), Bickerton (1990), and (Maynard Smith and Szathmáry, 1995, Ch. 17) have argued that human ability to acquire language is biologically based or innate. In particular features of the ambient language's grammar are acquired by setting parameters in a universal grammatical system for human language (Chomsky (1981)). This system might be inborn or developed, in that all humans acquire it in the course of development, and may have a large genetically transmitted component that is not merely part of general cognitive abilities and intelligence. Meanwhile, others argue that general human cognitive abilities will eventually be able to explain the origin and maintenance of language (e.g. Steels (1995)). Many workers are now studying the degrees to which innateness or competing mechanisms can serve as explanations of the evolution of linguisitic phenomena (e.g. Hurford et al. (1998)). One should resist the tendency to demonize generative grammar on the grounds that it seems to attribute discontinuity of capacity between humans and other animals. The emerging picture may be one in which human language acquisition has a strong evolutionary compotent with language specific developmental canalization that combines with more general aspects of cognition to generate language readiness (e.g. Batali (1994)). There is not enough evidence on either side to conclusively say that human language acquisition capacity is primarily innate or primarily based on culture and general cognitive abilities. Language readiness of humans may also have some unexpected sources, combining the evolution of neurophysiology with other abilities, e.g. see the discussion of mirror neurons in monkey brain area F5 (which fire both when particular affordances are used in action by the animal or observed being used in actions of others) which is homologous to Broca's area in human for a proposed model of human language evolution (Rizzolatti and Arbib (1998); Arbib (to appear)).

The degree to which communicative systems are innate, subject to developmental variation and learning, and whether their learned aspects are mainly acquired via individual or social learning are often topics of heated debate. Of course, the degree to which and which aspects of such systems are innate will vary considerably from species to species.

5 Shared Meaning

Having rejected privileged agent-independent notions of semantics, meaning, signs, concepts and mappings, how is it possible to account for the fact that agents do in fact succeed in cooperation and communication? Does this not require us to resort to postulating external Platonist universals to which agents have at least limited access? No, it does not.

Similarities of experience between agents begin to can account for the observed correspondences in the making of meaning. Agents sharing an environment, with similar sensory and actuator apparatus, with similar bodies and needs will to lesser or greater degrees share modes of interaction with their world. Their Umwelts (worlds of experience around the agent) may correspond to lesser or greater degrees. The sharing of these features can be the substrate supporting similarity of sensory perceptions, similarity of actions, and needs (hence of what is useful for the agents). This can already account for innate similarities in the experience of meaning, and hence of the grounding for communication via similaries between the sender and receiver (Dautenhahn (1995); Nehaniv et al. (1999)). However, the sender and receiver of a signal may have radically different embodiments, such as echo-locating bats and their insect prey, dolphins and prey fish. In such cases, the signal may also result in transfer of information useful to one or both parties, but the meaning of the interaction is only shared in the sense that both parties take part in two different instances of semiosis in which there is overlap in the signal and possibly the signified vertices of the semiotic triangle.

In societies of interacting agents, there is an opportunity not only for the signs and signified to converge within distinct agents, but building on biological factors, there is also opportunity for further convergence by means of learning in the course of many interactions. This may result in a convergence of concepts, signifieds, and conventionalizations of signals into systems of signs. Moreover, the mappings, linking signs and signifieds, may also have the opportunity to converge. In this case, shared semiotic systems make communication more like – but still very distinct from the Platonistic idealization and simplification of denotational semantics with its "sign-meaning pairs". Beyond biologically innate or developmentally 'programmed' instances of such convergence, conventionalization of interaction via cultural transmission or social learning appears to be the only possible mechanism that can account for the emergence of such (shared) semiotic systems.

In interspecies interaction, parrots (Pepperberg (to appear)), chimpanzees, bonobos (Savage-Rumbaugh and Brakke (1996)) and bottlenosed dolphins (Herman and Austad (1996)) have all shown that they are capable of acquiring various components of human or human-constructed language-like communication systems, involving categories and reference, requests to satisfy intentions, and in the case of bonobos and dolphins, also the ability to understand, as evidenced by action in controlled experiments, syntactically complex imperatives, or again for dolphins, even notions of absence and abstract concepts such as simultaneity (tandem action) and imitation (Herman (to appear)). Social interaction (with humans) was a key feature in the animals' acquisition of these linguistic abilities.

M. Oliphant (Oliphant (to appear)) argues that as far as we know only humans have naturally occurring arbitrary symbolic reference. He shows that learning such arbitrary correspondences (between "meaning-symbol" pairs) is easily accomplished already using very simple artificial neural network models (e.g. using Hebbian learning), so computational capacity limitations on learning ability cannot be responsible for the observed apparent lack of learned arbitrary referential symbols in non-human animals. He speculates that this lack may be due to the difficulty in "observing meaning", i.e. other animals do not learn to communicate because of difficulty in "determining the meaning a signal is intended to convey." Meanwhile, humans use taxonomic categories, awareness of pragmatic context, reading the intent of the speaker, and human adults modify their utterances when speaking to younger children.

However, experiments with socially-mediated learning in (even differently embodied) robotic agents, show that at least acquisition corresponding labelling ("proto-words") for similar external environments is possible via associative learning using temporal delays (Billard and Dautenhahn (1999)).

All of this suggests that shared meaning (corresponding processes of semiosis) requires shared experience in a social setting (or biologicial innate similarity). It is important in the social acquisition of sign systems that agents are allowed to attempt uses of communication to meet their own goals (e.g. intentions, homeostasis, transportation, feeding needs) rather than those of experimenter or other agents (Savage-Rumbaugh and Brakke (1996)). This is in accord with the notion that meaning depends on usefulness to the agents, and thus motivates the acquisition of the semiotic system, as when human children acquire human language.

6 Interaction Games

In this section, we will look inside communication and examine some of the most important features that are present in at least some forms of animal or human communication.

6.1 Language Games

Wittgenstein viewed natural language as comprised of myriad (and often very separate) language games in which language is employed in particular contexts by participants in particular manners. He constructed many examples of *language games* played according to strict rules in his philosophical investigations (Wittgenstein (1968)) to gain insight into the nature of language and other topics. In each game participants (or, agents, if you like) use language to accomplish certain things in the world. Wittgenstein uses the word 'grammar' to describe the use of language or language components (whether natural, formal, or artificial) in carrying out particular tasks or activities. Some examples of language games: children singing 'Ring around the rosy, a pocket full of posies, Ashes, Ashes, we all fall down' when dancing in a circle holding hands; making a list of items to buy at a grocery store, and then checking them off the list as they are collected into one's shopping basket; asking another person the time; yelling 'brick' or 'slab' at a construction site when asking another worker to bring the needed object. Many of Wittgenstein's examples include simple finite languages with strict rules of use, but the notion includes all ways in which natural language is employed.

Context is crucial in language games. When the rules of one game are applied in the context of another situation, interaction may fail, or we may produce in ourselves a sense of confusion or bewilderment. For example, the syntax of natural language allows us to say "Where is the book?", an ordinary question we might ask in trying to obtain an item. Since "the book" is a noun phrase, we might substitute another noun phrase such as "the universe" or "toothache" to create unusual questions, which seem meaningful since we can form them syntatically. Yet they are not part of our everyday life language games and so are not "grammatical" within these games. Similarly, since we can say "What happened before Thursday?", syntax allows us to say "What happened before time?". Much of philosophy begins with attempts to interpret such use of language outside the ordinary contexts of its uses in natural language games.

Real agents only play the language games that are useful for them. A statement like "This pen is blue" is never made about a pen that the speaker knows is red, unless there is a reason behind this. Examples games in which this could occur: the speaker wishes to deceive or manipulate others; the speaker is illustrating the possibility of counterfactuals (in doing philosophy - i.e. playing a philosophy language game).

6.2 Interaction Games

Generalizing Wittgenstein's notion of language games to non-linguistic realms, the author has described *interaction games* in which agents employ channels of sensing and actuation in some manner that is useful for them (Nehaniv (1999)). This is essentially the notion of language game, except that it has been minimally expanded so that it now easily applies to non-human animals and (robotic or software) agents. The notion of interaction games, includes animal communication and signalling (see below), and since the notion "useful" can be well-defined in terms of reproductive success or evolutionary terms, the identification and study of interaction games in the animal world provides part of the basis for understanding evolutionary continuity between humans and other animals. Such parsimony between explanations of human and animal features of interaction and communication is a major theme of evolutionary psychology (Byrne and Whiten (1988)), cognitive ethology (Griffin, 1976, p. 102), or the study of animal minds (Griffin (1992); Jamieson and Bekoff (1996)).

6.3 Games Animals Play

Formalized signalling interactions are apparent in the natural behavior of many animals. In dogs a 'play bow' may precede what would otherwise appear to be aggressive or sexual behaviour (Bekoff (1977)). Marking a sequence by a preceding play bow tells the canid observer "what follows is play". Squids, cuttlefish and octopi employ elaborate signalling systems for attracting a mate, threatening rivals, hunting, confusing or frightening others and for camouflage. Chromatophores in the skin of many cephalopods allow them via fast neural control to alter their body patterning, to signal to conspecifics or members of other species, even sending different signals to different observers viewing the animals from various perspectives (Moynihan (1985); Hanlon and Messenger (1996)). Squids can very quickly change from one display to another in a sequence. It is unclear whether and to what degree these changes are syntactically governed.

6.4 Comprehension / Production

Humans (and other animals or agents) may have differing capacity in comprehending as compared to producing communicative signals. Generally, but not always, ability to receive and interpret (parse or act on) communication is higher than the ability to produce communicative signals as evidenced in humans, apes, and dolphins (Herman and Austad (1996)).

6.5 Deixis

The indication of direction or directional reference to objects in language and interaction is called *deixis*. We see it in humans in deictic gaze (already present in prelinguistic infants) and also in words like "this" and "those".

Ants pheromones seem to have deictic qualities. And the use of honeybee dances to point in a sophisticated way that indicates both direction and distance is another example. Despite what is sometimes asserted, the honeybees' dances do not refer only to sources of food, but may be employed also for other deictic functions such as the indication of desirable nesting sites (Griffin (1976)).

6.6 Reference, Categories, and Naming

Labelling particular objects, or categories of objects is a property of human language. More generally, not only objects, but attributes, actions, and relationships can be named with words. Categories can group together entities based on functional similarity, i.e. the fact that they require similar behavioural responses, or on syntactic similarly, i.e. a degree of interchangability between words of the same category in the structure of utterances (e.g. transitive verbs, animate singular nouns, etc.) How such categories might arise in humans and animals is unclear. But artificial neural network models in which the output is to behavioural selection rather than classification might lead to insight. Clustering into categories can thus arise via separability, or via association of objects with similar properties (i.e. similar to the agent perceiving them).

Reference for proper names (signals labelling unique items, places or individuals) is less of problem than is the origin of abstract nouns, classes, categories, verbals, attributes, and relation words.

6.7 Association vs. Predication

Hebbian learning and concept formation using artificial neural networks may be sufficient for describing the phenomenon of association, and even for some cases of action selection. Association is generally symmetric, but can be made asymmetric, e.g. through the use of temporal delay information. *Predication* is a particular type of asymmetric association, e.g. "This pen is red" predicates a property of an entity ('pen' and 'red' are not merely associated). Predication, as in assertions that some entity has a property, has a weaker cousin *modification*, which is a function of adjectives and adverbs, which are responsible for a kind of less marked, presuppositional, predication in language. Grades of abstractness in predication depend on the notion of category (e.g. entity with proper name or generic entity) and attributes (properties). There seems to be no evidence for natural occurring instances of predication in non-human animals. Why this is so remains to be explained. Predication may lie at the core of human syntax. Another weaker version of it seen in human language includes topic comment constructions.

6.8 Discrimination Games

Pepperberg (to appear) presents evidence for predication, attribution of properties to objects, in African Grey parrots trained using a socially-based model rival technique. Apes can be taught to use attribute labels (e.g. Savage-Rumbaugh and Brakke (1996)) and bottlenosed dolphins demonstrate understanding of absence vs. presence of objects and distinguish possible vs. impossible requests in a syntactic command language used with them by human trainers (Herman and Austad (1996)). We can call games in which an agent must indicate or possibly even predicate that an entity has a property *discrimination games*. In many cases it is still unclear to what degree what is happening is like predication in human language.

This sort of interaction game is employed by Steels (1995) in experiments with software agents and robots. With possible referents given a priori in his model along with separation of sensory channels, individuals in the game attempt to refer to the same object in the environment. This goal of reference is built in, as is the notion of predication. Success in this game occurs if the predicate (given by the sender) uniquely determines the entity of which the predication is made to the recipient or determines a third entity whose spatial relationship to another has been predicated. Iterated playing of the the game leads to convergence of (proper) names labelling of entities, and of either spatial predicates that determine a third entity, or, alternatively, of predicates that constrain ranges of (sometimes several) feature values. Within each agent, phonetic symbols are associated to ranges of values in sensory channels. Communicative success is the criterion each agent uses in deciding whether or not to revise its association of phonetic elements to labels for objects or for attributes. Although the models of (Steels (1995)) have built in capacities for reference and predication, the system does illustrate how conventions of labelling can arise in a population that has such capacities, even if the set of objects and attributes is open and changing.

Explaining how reference and predication could arise remains an open problem.

6.9 Following Games

In *following games* (employing learning by imitation), signals are employed to ensure the coordinated movement of teacher and student robots. Additionally, short binary string signals ('words') are emitted by the teacher as a function of its sensor values. By using an appropriate delay parameter (related to body length and speed of motion), the student comes associate the words with its own sensory experience in similar contexts. Thus the 'meaning' of the signals is acquired (Billard and Dautenhahn (1999)). Here the signals are from a small finite set, but the perceptions they are associated with need not be similar since the technique works even with agents having different body architectures.

7 Syntax

Syntax (rules of grammar) is often considered by linguists as being absolutely necessary for human like linguistic ability. Some precursors and features are the combination of symbols to yield new types of communicative acts not previously possible (Savage-Rumbaugh and Brakke (1996)), rule sets generating finite sets of possible signalling events, compositional or subcategorization structure, and recursion and combinatorial explosion in the number of possible communicative acts (see below).

7.1 Compositional Structure

The language used by Herman and Austad (1996) with dolphins had a strict word order in which target goals occur first, objects to be manipulated occur in second position, and actions occur last. While still finite (though extensible), this language has *compositional syntactic structure*: commands in the language take arguments whose role is determined by position. Allowing other marking (other than position) to indicate role would also yield compositional syntax.

Lexical items can take arguments (subcategorization), e.g. VP $\rightarrow$ V NP, a verb phrase may be constituted from a verb followed by a noun phrase as in $[_{VP}[$ eats $]_V[_{NP}$ the chocolate cake $]_{NP}]_{VP}$. Grammatically "the chocolate cake" is the direct object of the "eats". "Eats" has constituents or slots, including an object slot. The correspondence between the argument structure in syntax and semantics is also sometimes called 'compositionality' (e.g. Kirby (1999)), but this might more precisely be called homomorphic mapping or morphism or, more generally, a structure-preserving map (e.g. Goguen (1999)), i.e. the terms of logical form, syntatic representation, and phonetic form can be obtained via structure preserving correspondences. This is what Chomsky calls the 'projection principle' (Chomsky (1981),(Sells, 1985, p. 33)).

7.2 Recursion

When a lexical item subcategorizes for other items, it may be that by following a chain of such subcategorizations that it is possible to reach an another item of the original type. E.g. "I believe that you think ...", in such cases recursion is possible. Or in phrase structure rules

$$X \stackrel{*}{\rightarrow} \alpha X \beta,$$

where X is a non-terminal and α, β are some strings. More generally, exponential growth in the number of generated strings can result when there are derivations of the form

$$X \stackrel{*}{\rightarrow} \alpha X \beta$$

with α and β non-empty or

$$X \stackrel{*}{\rightarrow} \alpha X \beta X \gamma.$$

Recursion and related exponential growth in generative capacity are extremely likely to arise in random sets of rules for context-free grammars.

Formal language theory, concerned with the description of sets of strings, provides convenient methods to describe such structure. Chomsky's *Syntactic Structures* (Chomsky (1957)) shows that (while English is not a context- free language) a context-free formal grammar can give an approximation of a fragment of English. The same holds for other human natural languages. The formalism works well for computer languages such as PASCAL, FORTRAN, C, etc., which are actually defined using such formalisms. Semantics of these languages is compositional in the sense that fixed meanings percolate up from leaf nodes in the parse tree of the language statement, and functions at intermediate nodes are applied to the node's constitutent argument list. (E.g., consider how an assigment statement like $X := C + 5$ is parsed: the value of variable "C" and integer "5" are arguments to function "+", so that $C + 5$ comprises an expression evaluated by applying addition to these arguments; while the assignment operator ":=" takes a variable and expression as its arguments, evaluates the expression and assigns the result to the variable X.).

First-order and higher-order logic formulae are similarly constructed using context-free grammars. Truth values of formulae in the languages determined by these grammars are similarly determined (with respect to a particular structure or "world of discourse" over which the interpretation is made) by recursive application of rules which finally reduce to the assignment of truth values to the equality of terms and the truth values of predicates. Rather than inducing well-defined operations in a computer, the interpretation of a logical formula over a structure returns either "true" or "false". Once the structure and rules of interpretation have been thus specified, all observers will assign the same truth value to each formula.

Predication is built into the ediface of formal logic. Constituent argument structure ("compositionality") is built into the formalism of first-order logic and into the structure of programming languages, and other formalisms. These properties were abstracted from natural language by logicians and mathematicians. They have been codified and standardized in such a way that someone using them is able to 'escape from context', i.e. knowledge expressed in such formulae is an example of what Bruno Latour (Latour (1987) has called an 'immutable mobile', knowledge that can be reused in other contexts when applied according to certain general procedures or rules. Joesph Goguen has called this 'dry' information, as opposed to 'wet' information which cannot be interpreted outside its particular original grounded, embedded, situated context. Note that there are degrees of dryness and wetness, or in Latour's terms, degrees of mobility. For example, a cake recipe, is a partly formal but reusable piece of information somewhere in the middle of the wet-dry continuum.

These formal properites of compositionality (argument structure, subcategorization) and semantics of predication are thus very well-supported by the tools of computer science and formal grammars. It is very easy to describe compositional formal language systems and associated semantics using these tools. That is exactly what the tools were developed for. Tools such as context-free grammar (Backus-Naur form), phrase structure grammars, denotational semantics, programming languages, etc., abstract from structure of natural human language and also 'clean-up' the embeddedness ('wetness') increasing the mobility of knowledge (well-definedness of truth values of formulae when interpreted over structures, portability of software, etc.)

It should therefore come as no surprise if we observe the "emergence" of predication or compositionality or of recursion in models of the evolution of communication and evolution of language which formulate their grammars using tools of context-free grammar or subcategorization in argument structure: The latter were constructed to facilitate the former.

8 Random drift: "Diversity" and "Convergence"

In repeated stochastic sampling of a population, the distribution in the sample is unlikely to exactly match the distribution of characters in the population. This phenomenon is well-known in statistics, where large sample sizes and confidence intervals are used to limit and quantify the likely effects of sampling error (Freedman et al. (1997)). In evolutionary genetics (Maynard Smith (1989); Sigmund (1993); Roughgarden (1996); Schmitt and Nehaniv (1999)) repeated sampling of a finite population (and all biological populations are finite) results in genetic drift of the inherited traits (independent of natural selection and variation due to mutation) towards random but uniform values. Explicit bounds on the rate of convergence due to genetic drift in iterated random sampling with or without the action of selective pressure have been calculated (see the above references). It is a mathematical theorem, that under very general circumstances, e.g. in the absence of mutation, a fixed-size finite population subject to any operators of fitness selection and with or without sexual recombination will converge (with probability 1) to a population of individuals all having the same genotype. Moreover, this is even true, if for instance, what is transmitted is not called 'genotype' but is e.g. a 'meaning-symbol' map acquired from observation of other agents' use of 'language'. This is all that is behind the so-called 'emergence of a common language' in some computational models. Sometimes such random drift convergence has been given the name 'self-organization'.

Convergence can be prevented by the introduction of random variation in the course of reproduction (e.g. the random resetting of bits in a genetic algorithm). These mechanisms by themselves explain much of what is seen e.g. in the results of Arita and Koyama (1998) on so-called "linguistic diversity".

Cases of random drift and drift combined with selection and variation are seen, for example, in the studies of Arita and Koyama (1998) at a genetic level for individuals defined by meaning-symbol pairs, of Steels (1995) in which entities consist of sets of meaning-symbol pairs but modify themselves (selection and variation) based on communicative success, and of (Hashimoto and Ikegami (1995); Steels (1998); Kirby (1999)) in which individuals can at least roughly be viewed as grammars, i.e. populations of sets of rules.

'Emergence' and 'self-organization' are terms used by experimenters to describe phenomena which surprise them and for which they can offer no detailed explanation. Minsky has argued that use of the word 'emergence' should make one suspicious that not enough effort has been made in finding explanatory mechanisms (Minsky (1996)). If the criterion for emergence is one of surprising the investigators, then the notion is clearly very much observer-dependent, in such a formulation of little value to science. However, emergence can be defined in a more formal way in terms of a rigorous mathematical definition of complexity as complexity increase in the extreme upper range of certain bounds on complexity increase (for one-way interactions) or greater increase (for interaction with feedback), see (Nehaniv and Rhodes (2000)).

9 Building the Solutions In

We have seen some evidence that simulation models without evolution of innate language ability can be put forward for to assess possible explanatory mechanisms for aspects of language or communication evolution. Steels' discrimination games (Steels (1995)) have also been extended to games in which not only phonemic labels, but constraints on ordering are introduced to model evolution of syntax (Steels (1998)). In the former predication and reference were built in to the agents, in the latter subcategorization frames are built in, i.e. compositionality is assumed, although not its particular realization under a mapping to 'surface structure'. Kirby (1999) starts with a space of privileged meanings that are compositional and recursive, and using context-free formalisms to acquire grammars which define structure-preserving maps from 'meanings' to 'utterances'; in this setting he shows that the bottleneck of learning (and certain generalizing variation operations) leads over time to increasingly generic context-free grammars that preserve structure of the external 'meaning' space. Hashimoto and Ikegami (1995) show that social factors can determine the communicative success of grammar using agents that play a game of generating and parsing abstract utterances. Subjacency, a structural constraint on argument chains in determining reference in universal grammar (e.g. (Sells, 1985, p. 48)) can probably be shown to arise once context-free like rules are employed in compositional syntax. The origin and maintenance of syntactic phenomena such as deixis, predication, compositionality, and grammars can still be considered wide open problems.

Innate language acquisition devices and language readiness (either neurophysiological, cognitive, or cultural) have been proposed but yet not demonstrated as sufficient to account for human linguistic capacities (Chomsky (1968); Pinker and Bloom (1990); Arbib (to appear); Hurford et al. (1998)). We expect a crucial role for social factors and interaction, at the level of individual development and in evolving populations or societies of agents.

References

Michael A. Arbib. The mirror system, imitation, and the evolution of language. In K. Dautenhahn and C. L. Nehaniv, editors, *Imitation in Animals and Artifacts*. MIT Press, to appear.

Takaya Arita and Yuhiji Koyama. Evolution of lingusitic diversity in a simple communication system. *Artificial Life*, 4(1):109–124, 1998.

J. Batali. Innate biases and critical periods. In Rodney Brooks and Patti Maes, editors, *Artificial Life IV*, pages 160–171. MIT Press, 1994.

Marc Bekoff. Social communication in canids: Evidence for the evolution of stereotyped mammalian display. *Science*, 197:1097–1099, 1977.

D. Bickerton. *Language and Species*. Chicago, 1990.

A. Billard and K. Dautenhahn. Experiments in learning by imitation - grounding and use of communication in robotic agents. *Adaptive Behavior*, 7 (3/4), 1999.

Jack W. Bradbury and Sandra L. Vehrencamp. *Principles of Animal Communication*. Sinauer, 1998.

J. Bruner. The narrative construction of reality. *Critical Inquiry*, 18(1):1–21, 1991.

Gordon M. Burghardt. Ontogeny of commuication. In Thomas A. Sebeok, editor, *How Animals Communicate*. Indiana University Press, 1977.

Richard Byrne and Andrew Whiten, editors. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford, 1988.

Noam Chomsky. *Syntactic Structures*. Mouton Publishers, 1957.

Noam Chomsky. *Language and Mind*. Harcourt, Brace and World, 1968.

Noam Chomsky. *Reflections on Language*. Pantheon Books, 1975.

Noam Chomsky. *Lectures on Government and Binding*. Foris Publications, 1981.

Charles Darwin. *The Expression of the Emotions in Man and Animals*. 1872.

Kerstin Dautenhahn. Getting to know each other – artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356, 1995.

David Freedman, Robert Pisani, and Roger Purves. *Statistics*. W.W. Norton & Co., 3rd edition, 1997.

Joseph A. Goguen. An introduction to algebraic semiotics, with application to user interface design. In *Computation for Metaphors, Analogy, and Agents*, volume 1562, Lecture Notes in Artificial Intelligence, pages 242–291, 1999.

Donald R. Griffin. *The Question of Animal Awareness: Evolutionary Continuity of Mental Experience*. The Rockefeller University Press, 1976.

Donald R. Griffin. *Animal Minds*. Chicago, 1992.

Roger T. Hanlon and John B. Messenger. *Cephalopod Behaviour*. Cambridge University Press, 1996.

Takashi Hashimoto and Takashi Ikegami. Evolution of symbolic grammar system. In F. Morán, A. Moreno, J. J. Merolo, and P. Chacón, editors, *Advances in Artificial Life (3rd European Conference on Artificial Life, Granada, Spain, June 1995)*, volume Lecture Notes in Artificial Intelligence, 929, 1995.

Louis M. Herman. Vocal, social, and self imitation by bottlenosed dolphins. In K. Dautenhahn and C. L. Nehaniv, editors, *Imitation in Animals and Artifacts*. MIT Press, to appear.

Louis M. Herman and Steven N. Austad. Knowledge acquisition and asymmetry between language comprehension and production: Dolphins and apes as general models for animals. In Marc Bekoff and Dale Jamieson, editors, *Readings in Animal Cognition*, pages 289–306. MIT Press, 1996.

James Hurford, Chris Knight, and Miceal studdert Kennedy, editors. *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press, 1998.

Dale Jamieson and Marc Bekoff. On the aims and methods of cognitive ethology. In Marc Bekoff and Dale Jamieson, editors, *Readings in Animal Cognition*, pages 65–77. MIT Press, 1996.

Simon Kirby. Learning, bottlenecks, and infinity: a working model of the evolution of syntactic communication. In K. Dautenhahn and C. L. Nehaniv, editors, *Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts*, pages 55–63. Society of the Study of Artificial Intelligence and the Simulation of Behaviour, 1999.

Bruno Latour. *Science in Action*. Harvard, 1987.

Bruce MacLennan. Synthetic ethology: An approach to the study of communication. In *Artificial Life II*. Addison Wesley, 1992.

John Maynard Smith. *Evolutionary Genetics*. Oxford University Press, 1989.

John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. W.H. Freeman, 1995.

Marvin Minsky. Plenary address at Artificial Life V, the 5th international workshop on the synthesis and simulation of living systems. 1996.

Martin Moynihan. *Communication and Non-Communication by Cephalopods*. Indiana University Press, 1985.

Chrystopher L. Nehaniv. Meaning for observers and agents. In *IEEE International Symposium on Intelligent Control/Intelligent Systems and Semiotics, ISIC/ISAS'99 (September 15-17, 1999 Cambridge, Massachusetts)*, pages 435–440, 1999.

C. L. Nehaniv, K. Dautenhahn, and M. J. Loomes. Constructive biology and approaches to temporal grounding in post-reactive robotics. In *Sensor Fusion and Decentralized Control in Robotics Systems II (September 19-20, 1999, Boston, Massachusetts), Proceedings of SPIE Vol. 3839*, pages 156–167, 1999.

Chrystopher L. Nehaniv and John L. Rhodes. The understanding and evolution of biological complexity from an algebraic perspective. *Artificial Life*, 6 (1), 2000.

Michael Oliphant. Rethinking the language bottleneck: Why don't animals learn to communicate? In K. Dautenhahn and C. L. Nehaniv, editors, *Imitation in Animals and Artifacts*. MIT Press, to appear.

Charles S. Peirce. *Collected Papers, Volume 2: Elements of Logic*. Harvard, 1995.

Irene M. Pepperberg. Allospecific referential speech acquisition in grey parrots (*Psittacus erithacus*): Evidence for multiple levels of avian vocal imitation. In K. Dautenhahn and C. L. Nehaniv, editors, *Imitation in Animals and Artifacts*. MIT Press, to appear.

S. Pinker and P. Bloom. Natural language and natural selection. *Brain & Behavioral Sciences*, 13: 707–784, 1990.

G. Rizzolatti and M. A. Arbib. Language within our grasp. *Trends in Neurosciences*, 21(5):188–194, 1998.

Jonathan Roughgarden. *Theory of Popultion Genetics and Evolutionary Ecology: An Introduction.* Prentice Hall, 1996.

Sue Savage-Rumbaugh and Karen E. Brakke. Animal language: Methodological and interpretative issues. In Marc Bekoff and Dale Jamieson, editors, *Readings in Animal Cognition*, pages 269–288. MIT Press, 1996.

L. M. Schmitt and C. L. Nehaniv. The linear geometry of genetic operators, with applications to the analysis of genetic drift and genetic algorithms using tournament selection. In *Mathematical and Computational Biology*. American Mathematical Society, 1999.

Peter Sells. *Lectures on Contemporary Syntactic Theories.* Center for the Study of Language and Information, Stanford, 1985.

R.M. Seyferth and D . L. Cheney. Vocal development in vervet monkeys. *Animal Beahviour*, 34:1640–1658, 1986.

C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, 1963.

Karl Sigmund. *Games of Life: Explorations in Ecology, Evolution and Behaviour.* Penguin, 1993.

W. John Smith. *The Behavior of Communicating: An Ethological Approach.* Harvard, 1977.

W. John Smith. Communication and expectations: A social process and the operations it depends upon and influences. In Marc Bekoff and Dale Jamieson, editors, *Readings in Animal Cognition*, pages 243–255. MIT Press, 1996.

Luc Steels. A self-organizing spatial vocabulary. *Artificial Life*, 2:315–332, 1995.

Luc Steels. The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103:1–24, 1998.

Ludwig Wittgenstein. *The Blue and Brown Books.* Harper & Brothers, 1958.

Ludwig Wittgenstein. *Philosophical Investigations (Philosophische Untersuchungen (German with English Translation).* Basil Blackwell, 3rd edition, 1968.